

## Environmental Data Discovery and Transformation

The Environmental Data Discovery and Transformation (EnDDaT) service is a tool used to discover data from a variety of data sources, collect and process the data, and perform basic transformations. The end result is that environmental data from multiple sources is sorted into a single table. This user guide will step through the process of obtaining data, and describe available transforms and processing available for the data.

## Contents

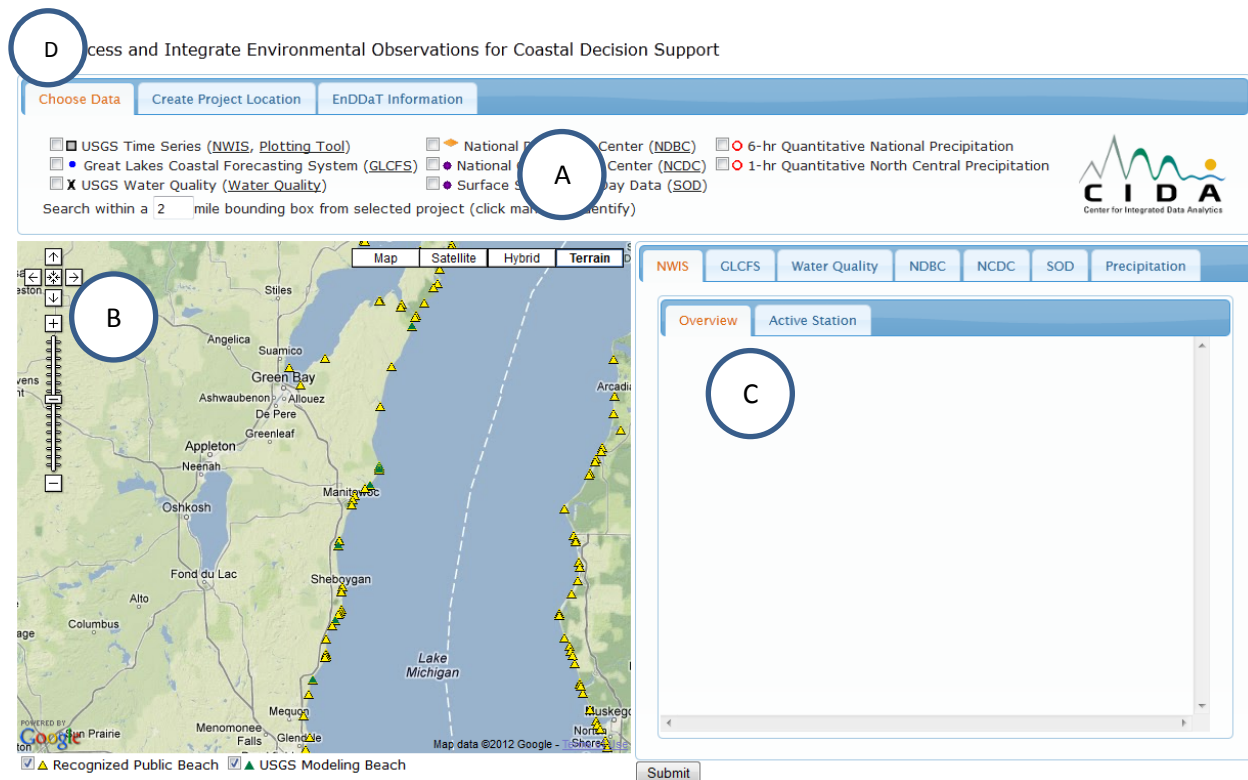
Environmental Data Discovery and Transformation.....	1
Overview .....	2
Data Discovery: .....	2
Select project location .....	2
Choose bounding box .....	3
Choose data to discover.....	3
View available data .....	4
Shapefile options .....	4
Select data and submit.....	5
Additional notes on individual data sources.....	5
Time Range and Filter Options.....	8
Review data choices.....	8
Process and Transformation Options .....	10
Process Options.....	10
Beach Orientation Calculator:.....	11
Final Output .....	13
A. Tab Delimited:.....	13
B. Comma: .....	13
C. Interactive: .....	14
Appendix 1 – Available Data Sources.....	15
Appendix 2 – Calculation documentation.....	19
Appendix 3: Version Updates.....	23

## Overview

EnDDaT opens to an introduction page that gives overview and motivation information. Aside from some technical details, links to the user guide and the main data discovery page are displayed. This guide will first step through the required steps for discovering data, choosing a time range, choosing data processing, and retrieving the data. At the end of each section, optional features are described.

## Data Discovery:

The data discovery page has 3 main sections. The top set of tabs (A), the lower left map (B), the lower right set of tabs (C). Additionally, (D) is where the loading indicator is located.



## Select project location

The first step to data discovery is to choose a location to begin the data search. Choose a pre-existing location on the map (triangle icons) or create a new project location. To create a marker, choose the 'Create Project Location' tab from the top set of tabs:

The screenshot shows the "Create Project Location" tab selected in the EnDDaT interface. The top bar has three tabs: "Choose Data", "Create Project Location", and "EnDDaT Information". Below the tabs, there are input fields for "Latitude" and "Longitude", a "My Location" link, and a "Create marker" button. A "OR" label is between the "Latitude" and "Longitude" fields. A "AND Marker Name:" label is to the right of the "Create marker" button. A "Lake:" dropdown menu is set to "Lake Michigan". A callout (D) points to a location pin icon.

There are several options for creating a project marker name. The first is to input a specific latitude and longitude in the corresponding text boxes. Alternatively, clicking on 'My Location' will automatically fill the latitude and longitude text boxes with an approximation of your current location. Another option is to place a project marker on the map. Clicking on the blue marker icon (circled above) allows you to place an active marker (red) on the map. This red marker can be dragged to adjust the position.

Once the position of the marker is in place, click on the 'Create marker' button. This will turn the icon into a red triangle, meaning it is an 'active marker', and data sources can be discovered around it. If a different project is selected at a later time, the custom project will remain as a green triangle as long as the page is not refreshed. The marker name is optional. Choosing a lake from the dropdown is only required if requesting Great Lakes Coastal Forecasting System data.

### Choose bounding box

The screenshot shows the 'Choose Data' tab with three sub-tabs: 'Choose Data', 'Create Project Location', and 'EnDDaT Information'. Below the tabs, there are several data sources with checkboxes:
 

- ☐ USGS Time Series ([NWIS](#), [Plotting Tool](#))
- ☐ Great Lakes Coastal Forecasting System ([GLCFS](#))
- ☒ USGS Water Quality ([QW](#))
- ☐ National Data Buoy Center ([NDBC](#))
- ☐ National Climate Data Center ([NCDC](#))
- ☐ Surface Summary of Day Data ([SOD](#))
- ☐ 6-hr Quantitative National Precipitation
- ☐ 1-hr Quantitative North Central Precipitation

 At the bottom, there is a text input field labeled 'Search within a' followed by a circled '3' and the text 'mile bounding box from selected project (click marker to identify)'.

Once a project location has been picked, the bounding box for searching for data must be defined. Input a distance to define a square bounding box. Data will be discovered within this distance from the active project marker. *Special note for IE users: hitting the 'Enter' button after modifying this field will not update the map markers. Click on any part of the page \*except\* the map to update this field.*

### Choose data to discover

This screenshot is identical to the one above, showing the 'Choose Data' tab with the same data source checkboxes and the 'Search within a 3 mile bounding box' field.

Data discovery within the bounding box can now be achieved by clicking on the check boxes next to the various data sources. The initial click of these checkboxes will send a query out to the data source, and retrieve information on data source locations and available data. Subsequently turning off and on the checked boxes will hide or show the icons on the map. If a new project marker or different bounding box is selected, new queries will be sent to any data source that is checked. The data is represented on the map by the icon that is next to the checkbox. Additionally on this menu, each of the data sources has a link (in parentheses) to the data provider. Note that showing and hiding the markers on the map will not affect what is selected in the data request tabs (next section). The selections will however be lost if the bounding box is changed or a new project is selected.

## View available data

Choose Data
Project Location
Create Project Location
EnDDaT Information

☒ USGS Time Series (NWIS, [Plotting Tool](#))
☒ Great Lakes Coastal Forecasting System (GLCFS)
☒ USGS Water Quality (QW)
☐ National Data Buoy Center (NDBC)
☐ National Climate Data Center (NCDC)
☐ Surface Summary of Day Data (SOD)
☐ 6-hr Quantitative National Precipitation
☐ 1-hr Quantitative North Central Precipitation

Search within a 4 mile bounding box from selected project (click marker to identify)

NWIS
GLCFS
QW
NDBC
NCDC
SOD
Precipitation

OverviewActive Station

Uncheck all: ☐

Station: 0408544206

Property	Begin Time	End Time
<input type="checkbox"/> Discharge (cfs) Daily Mean	2004-10-01	2007-05-07

Station: 0408544207

Property	Begin Time	End Time
<input type="checkbox"/> Discharge (cfs) Daily Mean	2004-10-01	2007-03-16

Station: 0408544209

Property	Begin Time	End Time
<input type="checkbox"/> Discharge (cfs) Daily Mean	2004-10-01	2006-11-30

Submit

Once data is discovered, it is displayed on the map. Additionally, a list of each of the data sources is shown in the 'Overview' tab within the individual data tabs on the right. Each data source has some sort of unique ID (in the above example, NWIS has a station number) which can be clicked to find out where they are on the map. Alternatively, an icon on the map can be clicked to get a more detailed description of the specific data source.

## Shapefile options

Map
Satellite
Hybrid
Terrain

NWIS
GLCFS
QW
NDBC
NCDC
SOD
Precipitation

Shapefile6 Hr National Points1 Hr Regional Points

Choose shapefile: fischer
Group results by: AREA
Upload shapefile
Upload Shapefile

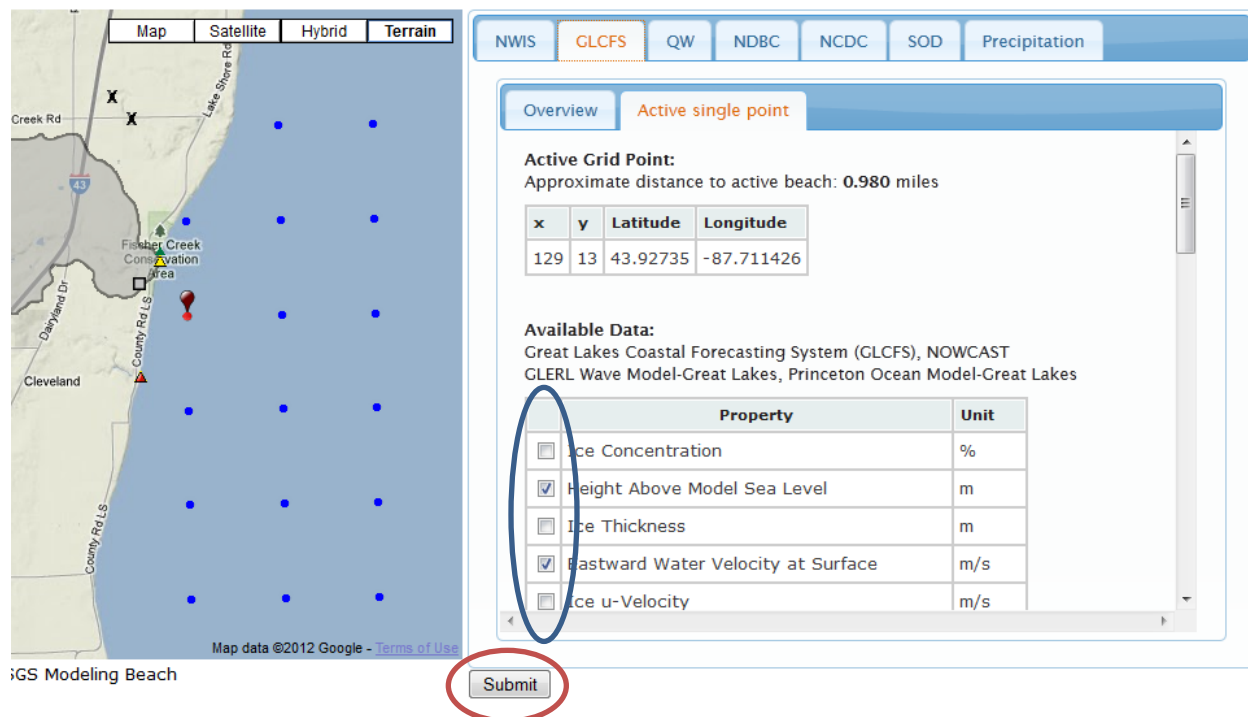
Shape file upload currently only works in Firefox and Chrome. Do not use any empty spaces in file name.
Choose one data set:
National Mosaic Quantitative Precipitation Archive (QPE) - radar-indicated rain verified and corrected precipitation estimates

Property	Start Date	End Date
<input type="checkbox"/> Historical Precipitation National River Forecasting Center 6 hr accumulation Covers U.S.	2000-10-01	2011-07-01
<input type="checkbox"/> Real Time Precipitation National River Forecasting Center 6 hr accumulation	2012-05-10	2012-05-17

The user can also upload a shapefile or choose one that has already been uploaded. This will be used only for the radar-indicated rain data. Select a zipped file containing (at a minimum) the .shp, .shx, .prj,

and .dbf components of your shapefile. Do not use empty spaces in the file name. The shapefile uploading tool currently does not work in Internet Explorer browsers. The user must choose a geographical processing to perform over the shapefile (mean, minimum, maximum, variance, standard deviation, weighted sum, or count). Explanations of these processes can be found in Appendix 2.C.

## Select data and submit



The screenshot shows the iGS Modeling Beach web application. On the left is a map of Lake Erie with a grid of blue dots representing data points. A red pin is placed on one of the dots. The map includes labels for 'Creek Rd', 'Lake Shore Rd', 'Fisher Creek Conservation Area', 'Dayland Dr', 'Cleveland', and 'County Rd LS'. At the bottom of the map, it says 'Map data ©2012 Google - Terms of Use' and 'iGS Modeling Beach'. On the right is a control panel with tabs for 'NWIS', 'GLCFS', 'QW', 'NDBC', 'NCDC', 'SOD', and 'Precipitation'. The 'GLCFS' tab is selected. Below the tabs are two sub-tabs: 'Overview' and 'Active single point'. The 'Active single point' sub-tab is active, showing 'Active Grid Point: Approximate distance to active beach: 0.980 miles'. Below this is a table with columns 'x', 'y', 'Latitude', and 'Longitude', containing the values '129', '13', '43.92735', and '-87.711426'. Underneath is a section titled 'Available Data:' listing 'Great Lakes Coastal Forecasting System (GLCFS), NOWCAST' and 'GLERL Wave Model-Great Lakes, Princeton Ocean Model-Great Lakes'. A table follows with columns 'Property' and 'Unit'. The table lists five properties with checkboxes: 'Ice Concentration' (unchecked, unit '%'), 'Height Above Model Sea Level' (checked, unit 'm'), 'Ice Thickness' (unchecked, unit 'm'), 'Eastward Water Velocity at Surface' (checked, unit 'm/s'), and 'Ice u-Velocity' (unchecked, unit 'm/s'). A blue circle highlights the checkboxes for 'Height Above Model Sea Level' and 'Eastward Water Velocity at Surface'. At the bottom of the panel is a 'Submit' button, which is circled in red.

x	y	Latitude	Longitude
129	13	43.92735	-87.711426

Property	Unit
<input type="checkbox"/> Ice Concentration	%
<input checked="" type="checkbox"/> Height Above Model Sea Level	m
<input type="checkbox"/> Ice Thickness	m
<input checked="" type="checkbox"/> Eastward Water Velocity at Surface	m/s
<input type="checkbox"/> Ice u-Velocity	m/s

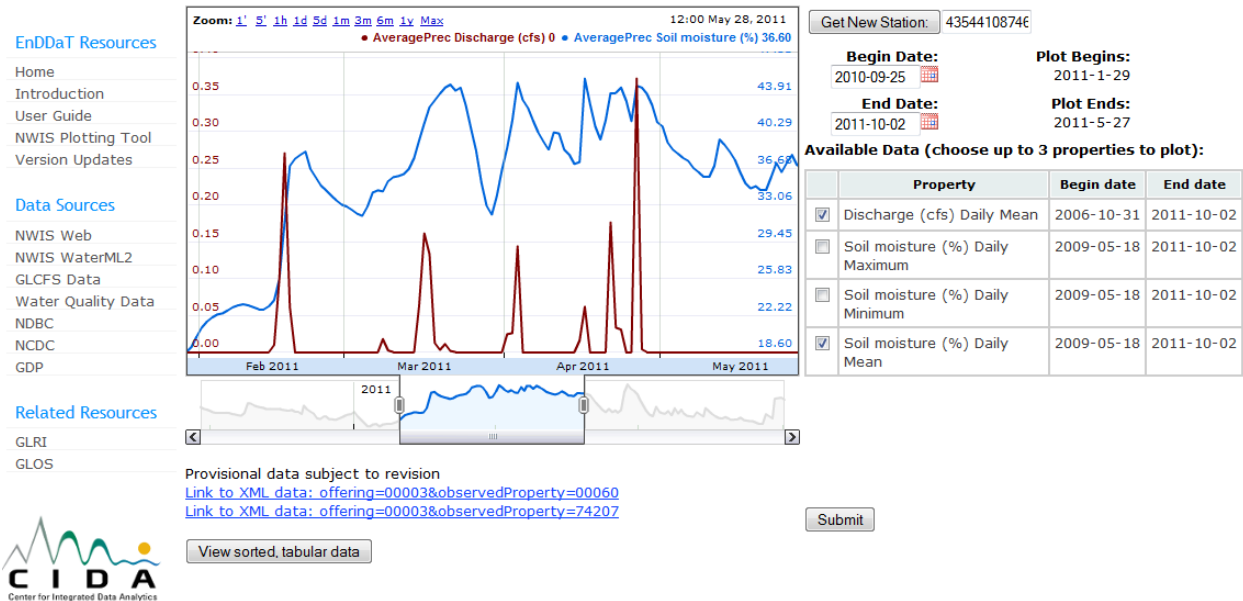
Using the check boxes, choose data to retrieve. When all the required data is requested, click the 'Submit' button circled in red. This will take you to the next page.

## Additional notes on individual data sources

### National Water Information System - NWIS

A complete list of available properties and corresponding time ranges is available on the Overview tab. Mousing over the station numbers will indicate where the stations are on the map. Clicking on those station numbers will move the active map marker to that station and load available properties to the Active Station tab. Alternatively, clicking on a NWIS marker on the map will load the selected available properties in the Active Station tab. Additionally, EnDDaT will check for data in the Instantaneous Data Achieve.

Clicking on a property will open a new window to the NWIS plotting tool. A graph of the selected property will appear. Any available NWIS properties can be plotted on this page. This page can also be accessed from the sidebar. Up to three properties can be plotted at once.



To change the plot, click on the requested properties from the 'Available Data' table on the right (up to three). Choose a date range, and click the submit button. Zoom and pan features are available on the plot either by using the small bottom plot, or using the mouse. Clicking on the plot, the scrolling with the middle scroll button will zoom, and using the left mouse button will pan. Links to the raw data are provided below the plot. To get data for a new station, type the station number in the upper-right 'Get New Station' text box, then click the 'Get New Station' button.

## Great Lakes Coastal Forecasting System (GLCFS)

A list of available grid points is initially displayed in the Overview tab. If a user wants to select multiple grid points, they must remember to select the points from this list, and then scroll down to also select the requested properties. Alternatively, a grid point can be selected (either from the map or the list). The available properties are then displayed in the 'Active single point' tab. Be aware that if properties are selected in this tab, then another grid point is selected, the selected properties are not saved and must be chosen again.

## Water Quality

A list of available stations, along with station identifier and type is initially displayed in the Overview tab. Clicking on a station name or a marker on the map will populate the available properties in the Active Stations tab. This tab is a little different than most of the other data sources in that it continues to populate as more stations are chosen. If the list becomes too large, the user can click the 'Clear list' check box, and start over. This was done because it is not feasible to list all of the available properties in the Overview tab (the query would take too long to load).

## **National Data Buoy Center (NDBC)**

A list of buoys is populated in the Overview tab. Clicking on a NDBC map marker or the buoy name in the table will populate the available properties in the Active Station tab.

*Note: the data has been temporarily disabled to fix a bug. We hope to get this back running soon. We also plan to modify the data discovery to be similar to the QW process, where clicking on individual buoys continues to populate an Active Stations tab, to allow for multiple buoys to be requested.*

## **National Climate Data Center (NCDC) and Surface Summary of the Day (SOD)**

Data discovery for the National Climate Data Center and Surface Summary of the Day services are offered. However, the data is not actually provided as output for EnDDaT. Both services require subscriptions to access the data. Links to the data are provided. We will look into ways to incorporate this data in the future.

## **Precipitation**

Total precipitation within a shapefile was discussed above. Additionally, precipitation from the individual grid points can be retrieved. Since there is only one property associated with a grid point, a list of available grid points is populated when the data source is selected (that is, there is no need for an 'Active Grid point' tab). When a grid point is chosen on the map or from the list, it will be indicated at the top of the appropriate tab.

## Time Range and Filter Options

### Review data choices

**A** **Start Date:** 2011-09-02 **B** **End Date:** 2011-10-02 **I** **Process Data**

**C** **Output date format:** MM/DD/YYYY HH:MM **D** **Output time zone:** GMT

**E** **Upload Times** **F** **Acceptable data gap:** 6 hours **G** **Filter file ID:**

**Optional:** Upload a file with dates and times to filter output.  
File should be a text file containing a single column of dates in the form: mm/dd/yyyy hh:MM  
See the [User Guide](#) for detailed information.

### NWIS:

**H**

	Property Name	Site	Property	Offering	Start Date	End Date
<input checked="" type="checkbox"/>	Discharge (cfs) Daily Mean	435440087463800	00060	00003	2006-11-09	2011-10-02
<input checked="" type="checkbox"/>	Temperature, water (deg F) Daily Mean	435440087463800	00011	00003	2009-05-18	2011-10-02

### Instantaneous Data Archive:

No IDA data requested

### Water Quality:

	Site	Property	Start Date	End Date
<input checked="" type="checkbox"/>	USGS-040854395	Total Chemical oxygen demand, (low level) [mg/l]	2007-06-28	2007-06-28
<input checked="" type="checkbox"/>	USGS-040854395	Bed Sediment Phosphorus [mg/kg]	2007-06-28	2007-06-28

### Options:

- A. Choose a starting date for requested data
- B. Choose an ending date for requested data
- C. Choose the date formatting for output data
- D. Choose output time zone (XST = X Standard Time, XDT = X Daylight Savings Time). Be sure to choose this before uploading a timestamp filter file.
- E. Click this 'Upload Times' button to upload a text file with a list of timestamps to filter the output data. If a file is not uploaded, the entire time series will be retrieved.

Acceptable formats:

7/5/2010 13:30  
7/6/2010 09:30  
7/7/2010 15:30



7/8/2010 13:30  
7/9/2010 13:30

Or

07/05/2010 13:30  
07/06/2010 09:30  
07/07/2010 15:30  
07/08/2010 13:30  
07/09/2010 13:30

The timezone of these dates/times should match the output timezone (part D).

- F. Only if using a timestamp filter (G), the Acceptable data gap is the how far back the filter will look for available data for one data point. For example, if the uploaded time file specifies data on 07/09/2010 13:30, and the Acceptable data gap is 6 hours – only data from 07:30 to 13:30 will be included on that line.
- G. A Filter file ID is given upon a successful time filter text file upload. If the user wants to same timestamp filter for multiple data retrievals, this id can be specified. Note that in this beta service, there is no guarantee that the files will be available for long-term.
- H. Review data choices. The checkboxes can be turned off to reduce the requested data
- I. Click the Process Data button to proceed to the processing data page

## Process and Transformation Options

### Process Options

The screenshot shows the 'Beach Orientation Calculator' interface. At the top, a blue bar contains the title 'Beach Orientation Calculator' and a tab labeled 'Choose Data and Data Processes'. Below this, the interface is divided into several sections:

- Available Data:** A list of data sources including 'Temperature, water (deg F) Daily Maximum', 'Discharge (cfs) Daily Mean', etc. Below the list are buttons for 'Add Selected', 'Add All', and 'Refresh'. A blue circle 'A' is placed over the list.
- Requested Data:** An empty list with buttons for 'Remove Selected', 'Remove All', and 'Refresh'. A blue circle 'B' is placed over this section.
- Data Processing:** A section with a 'Process' table and an 'Over:' section. The 'Process' table has columns for 'Process' and 'Over:'. The 'Over:' section has radio buttons for 'Hours', 'Days', and 'Weeks', and a list of time intervals (1 hour, 2 hours, 6 hours, 12 hours, 24 hours). A blue circle 'D' is placed over the 'Over:' section.
- Resolve Velocity Vectors with Beach Orientation:** A section with 'Direction' and 'Orientation' sub-sections. The 'Direction' sub-section has a radio button for '90 deg' and a radio button for '45.000 deg'. The 'Orientation' sub-section has checkboxes for 'Parallel (||)' and 'Perpendicular (⊥)'. A blue circle 'E' is placed over the 'Orientation' sub-section.
- Output Style:** A section with a dropdown menu for 'Tab Delimited Text' and a checkbox for 'As Download'. A blue circle 'F' is placed over the 'As Download' checkbox.
- Missing Value Fill:** A section with a dropdown menu for 'Missing Value Fill' and a button for 'Get Data'. A blue circle 'G' is placed over the 'Get Data' button.
- Buttons:** At the bottom, there are buttons for 'Get Data' and 'Generate URL'. A blue circle 'H' is placed over the 'Get Data' button, and a blue circle 'I' is placed over the 'Generate URL' button.

Arrows indicate the flow of data from 'Available Data' to 'Requested Data' and from 'Data Processing' to 'Requested Data'.

This page offers many options:

- View available data – the data that was requested from previous pages is listed here. One option can be clicked, or several using the mouse + control button, then choose 'Add Selected' to move the choices to the Requested Data column (B). Alternatively, 'Add All' will move all of the choices. The data options will remain in the column, but only the options in the Requested Data column will be retrieved.
- View requested data – only the data in this column will be retrieved. This list is populated by moving the 'Available Data' options as described in A.
- There are two options for moving data from the 'Available Data' column to the 'Requested Data' column. Select one property with the mouse, or multiple properties using the mouse + control button. The buttons Add Selected, Add All and Remove Selected, Remove All perform the same function as the >, >>, <, and << respectively. The Refresh buttons clear any selections from the list.
- Choose statistical processing – temporal statistics that can be applied to the available data. The user can choose one or several fields in Available Data (A), then choose the process and time. The resulting option is populated directly in the Requested Data column (B). Alternatively, if nothing is chosen in Available Data (A), then all of the data will appear in the Requested Data

column. The time over which the processing is calculated can be toggled between hours, days, and weeks. See Appendix 2A for calculation details.

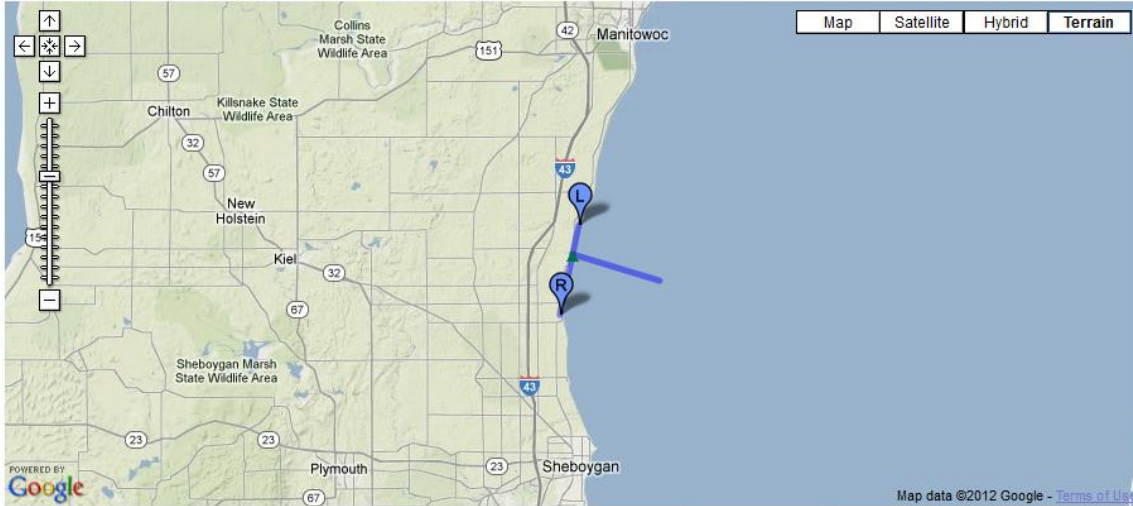
- E. Resolve velocity vectors – if there are sets of similar east and north vector components or wave direction (optionally paired with wave height), choosing parallel or perpendicular will resolve the vectors to the requested orientation. Choosing these will place an option in the Available Data box (A). The beach angle must be specified by the user, either in the text box, or using the beach angle calculator (see the next section Beach Orientation Calculator for more details). See Appendix 2B for calculation details.
- F. Choose output style. There are currently 3 options: tab delimited, comma delimited, and interactive. Interactive includes an interactive graph, as well as sortable table. There is also a checkbox here (Download) which allows the user to directly download the data.
- G. Choose a missing value fill. For instance, if there is no value – the user may wish to specify NaN. If left blank, the output is also blank.
- H. Get Data – this button is disabled until the data is ready to retrieve (currently, this is only an issue when gathering precipitation data from a shapefile). Once the data is ready, the button will be enabled, and the data gathering can be continued. If there was an error in gathering the precipitation data, a message will appear below the button. It is still possible to request the remaining data in that case.
- I. Generate URL – creates a URL of the requested data. Note that if a shapefile precipitation request is made, this URL is only valid for 7 days. All of the previous screens are only used to generate a URL for the final data retrieval. They could be circumvented by understanding how the URL is generated. This works especially well for running models on a daily basis. In a situation like that, we recommend going through the user interface once, then programmatically changing the beginPosition and endPosition. Alternatively, one could add 'latest=true' to get the latest data for the requested data.
- J. Beach Orientation Calculator tab. Choose this tab to bring up the beach orientation calculator

When all the requested data seems correct in the 'Requested Data' box, click the Get Data button (in the 'Choose Data and Data Processes' tab) to proceed.

### **Beach Orientation Calculator:**

If the beach orientation is unknown, it can be roughly calculated here. Align the L and R markers to be parallel to the beach, and the beach orientation angle will update in part E in the 'Choose Data and Data Processes' tab. The beach angle is calculated with north = 0°, and east = 90°. See Appendix 2B for details on the calculations.

## Calculate Beach Orientation



**Step 1:** Drag 'L' marker to left-most beach edge

**Step 2:** Drag 'R' marker right-most beach edge

**Step 3:** Perpendicular line should be pointed towards the water.  
If it is not, click the 'Flip Orientation' button

Flip Orientation

Marker	Latitude	Longitude
Left	43.9394	-87.7162
Right	43.8692	-87.7368

Beach orientation: **16.35°** from north

\* References to non-U.S. Department of the Interior (DOI) products do not constitute an endorsement by the DOI. By viewing the Google Maps API on this web site the user agrees to these [Terms of Service set forth by Google](#).

## Final Output

Screenshots of the output styles are shown here. If the 'As Download' button was checked on the 'Time Range and Filter Options' page, the output will be sent immediately as a download. It should also be noted that it is not required to go through the user interface to get this final output. A description of URL generation is listed in this section.

### A. Tab Delimited:

time	Height above sea level: [13,129]	Temperature, water (deg F) Daily Mean	Discharge (cfs) Daily Mean
09/01/2011 23:00	-0.011628735	NA	NA
09/02/2011 00:00	-0.0076217903	NA	NA
09/02/2011 01:00	9.15126E-4	NA	NA
09/02/2011 02:00	-0.0048775193	NA	NA
09/02/2011 03:00	-0.010676049	NA	NA
09/02/2011 04:00	-0.0039103352	NA	NA
09/02/2011 05:00	-0.01013631	NA	NA
09/02/2011 06:00	-0.008610567	NA	NA
09/02/2011 07:00	-0.010701612	NA	NA
09/02/2011 08:00	-0.003974536	NA	NA
09/02/2011 09:00	-0.024313048	NA	NA
09/02/2011 10:00	-0.008127013	NA	NA
09/02/2011 11:00	-0.015869694	NA	NA
09/02/2011 12:00	-0.013113555	68.0	0.000000
09/02/2011 13:00	0.00668042	NA	NA
09/02/2011 14:00	0.006515933	NA	NA

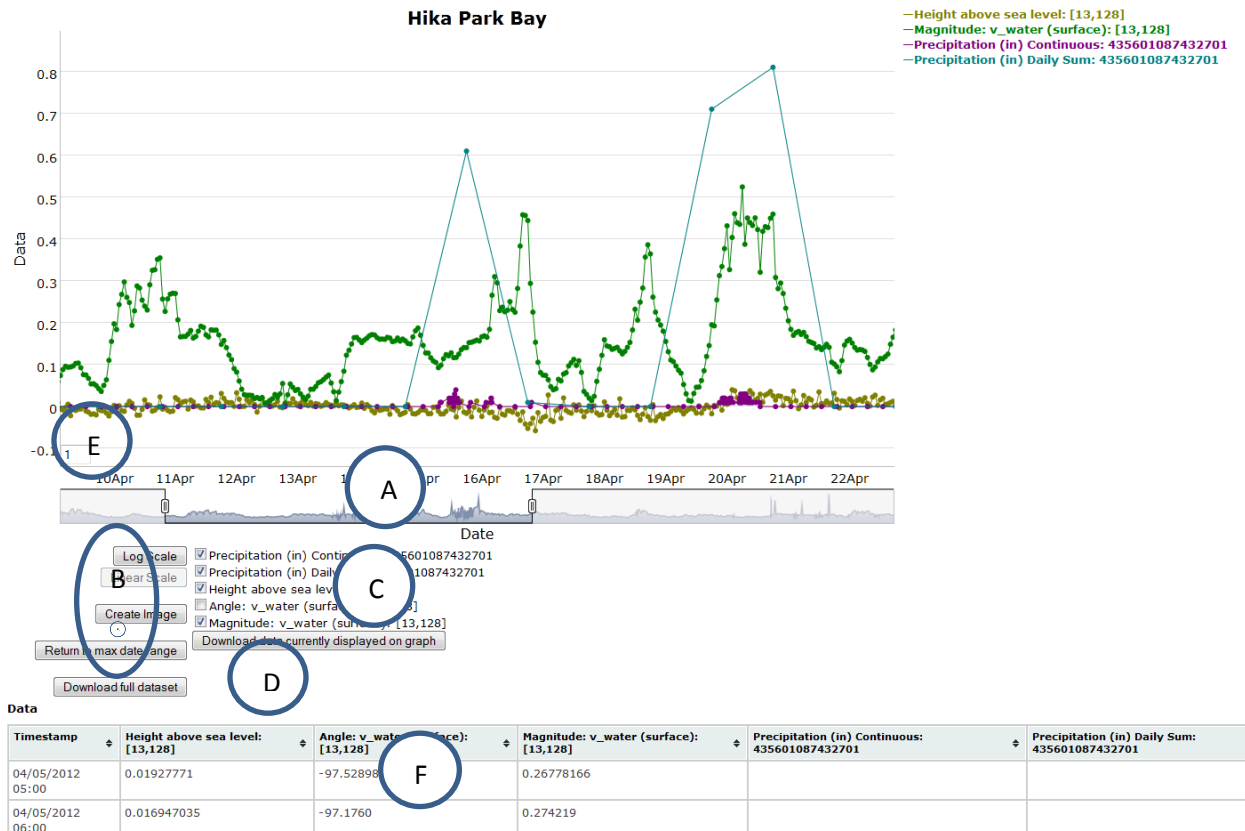
In this example, the Missing Value Fill from process data page was set to NA.

### B. Comma:

```
time,"Height above sea level: [13,129]","Temperature, water (deg F) Daily Mean",Discharge (cfs) Daily Mean
09/01/2011 23:00,-0.011628735,NA,NA
09/02/2011 00:00,-0.0076217903,NA,NA
09/02/2011 01:00,9.15126E-4,NA,NA
09/02/2011 02:00,-0.0048775193,NA,NA
09/02/2011 03:00,-0.010676049,NA,NA
09/02/2011 04:00,-0.0039103352,NA,NA
09/02/2011 05:00,-0.01013631,NA,NA
09/02/2011 06:00,-0.008610567,NA,NA
09/02/2011 07:00,-0.010701612,NA,NA
09/02/2011 08:00,-0.003974536,NA,NA
09/02/2011 09:00,-0.024313048,NA,NA
09/02/2011 10:00,-0.008127013,NA,NA
09/02/2011 11:00,-0.015869694,NA,NA
09/02/2011 12:00,-0.013113555,68.0,0.000000
09/02/2011 13:00,0.00668042,NA,NA
09/02/2011 14:00,0.006515933,NA,NA
```

In this example, the Missing Value Fill from process data page was set to NA.

### C. Interactive:



Options offered on this page:

- Zoom tool – slide the bars on the left and right to zoom in on the graph. Click within the zoomed section to pan left and right.
- Log/Linear scaling – toggle between a linear y axis and a log (base 10) y axis. ‘Return to max date range’ returns the graph to the maximum date range. ‘Create image’ freezes the graph in a .png image. From there, right click to ‘Save Image As’. ‘Download full data set’ – downloads all the data in tab delimited form.
- View data options – use these check boxes to show and hide individual data sets.
- Download data currently displayed on graph – downloads subset of data in tab delimited form.
- Smoothing – changing the number in this box adds a rolling average to the data.
- Sortable table – all the data is displayed in this table. Clicking on the arrows at the top of each column sorts the data based on that column.

## Appendix 1 – Available Data Sources

- **USGS Time Series (NWIS)**

National Water Information System

### Historical

Process	Property
Daily mean	Discharge/gage Height
Daily min	Water temperature
Daily max	Specific conductance
Daily sum	Dissolved oxygen
	pH
	Turbidity
	Precipitation
	Etc.
Continuous: (5-15 minute interval)	Discharge

Continuous – typically 15 minute data for past 120 days

Property
Discharge/gage height
Water Temperature
Specific Conductance
Dissolved Oxygen
pH
Turbidity
Etc.

Data Source:

<http://nwisvaws02.er.usgs.gov/ogc-swie/>

Data Type:

WaterML2 / SOS (OGC Standards)

- **NOAA Great Lakes Coastal Forecasting System (GLCFS), NOWCAST  
GLERL Wave Model-Great Lakes**

2006-Present

Hourly Data	3-Hour Data ( + choice of depth)
Height Above Model Sea Level	Sea Water Temperature
East/North Water Velocity at Surface	East/North Water Velocity
Depth-Averaged East/North Water Velocity	
Significant Wave Height	
Wave Direction, Period	
Air Temperature	
Cloud Cover	
Dew Point	
East/North Air Velocity	

**Data Source:**

<http://michigan.glin.net:8080/thredds/catalog.html>

**Data Type:**

Thredds/OpeNDaP

- **USGS Water Quality Data**

**Historic, irregular time steps**

Property	
Suspended solids	Phosphorus
Ammonia and ammonium	Nitrate-nitrite
Organic nitrogen	Oxygen
Hydrogen ion	Specific Conductance
Mercury	Etc.

**Data Source:**

<http://qwwbseervices.usgs.gov/portal.html>

**Data Type:**

WQX



- **National Data Buoy Center**

~2008-present, 30 minute data

Property	
Wind Speed, Direction, Vertical Velocity, Gusts	Water Temperature
Significant Wave Height	Wave Duration
Dominant Wave Period	Sampling Rate
Average Wave Period	Mean Wave Direction Peak Period
Swell Height, Period	Swell Wave Direction
Wind Wave Height	Wind Wave Direction
Wind Wave Period	Water Temperature

**Data Source:**

<http://sdf.ndbc.noaa.gov/sos/server.php>

<http://sdf.ndbc.noaa.gov/sos/>

**Data Type:**

SOS (OGC Standard)

- **Radar-indicated-rain**

**Historical Data:**

- 6 hour data
  - National RFC QPE Mosaic – covers country
  - 2000-10-01 – 2011-07-01
- 1 hour data
  - North Central River Forecasting Center – covers WI, MN, MI, IL, IN, parts of OH (So, Michigan, Superior, Huron, and parts of Erie)
  - 2000-06-15 – 2011-10-18

**Real Time:**

- 6 hour
  - National RFC QPE Mosaic – covers country
  - Latest week

**Data Source:**

<http://www.hpc.ncep.noaa.gov/npvu/>

**Data Type:**

Thredds

## Appendix 2 – Calculation documentation

### A. Statistical Processes

The following statistical processes are available: mean ( $\mu$ ), minimum, maximum, and summation( $\Sigma$ ), difference( $\Delta$ ), and standard deviation( $\sigma$ ). They require the user to specify a time period (e.g. mean over 6 hours). The equations are as follows:

$$\mu = \frac{\sum_{i=0}^N x_i}{N} \quad 1$$

$$\min = \text{minimum}[x_0, x_N] \quad 2$$

$$\max = \text{maximum}[x_0, x_N] \quad 3$$

$$\Sigma = \sum_{i=0}^N x_i \quad 4$$

$$\Delta = x_N - x_0 \quad 5$$

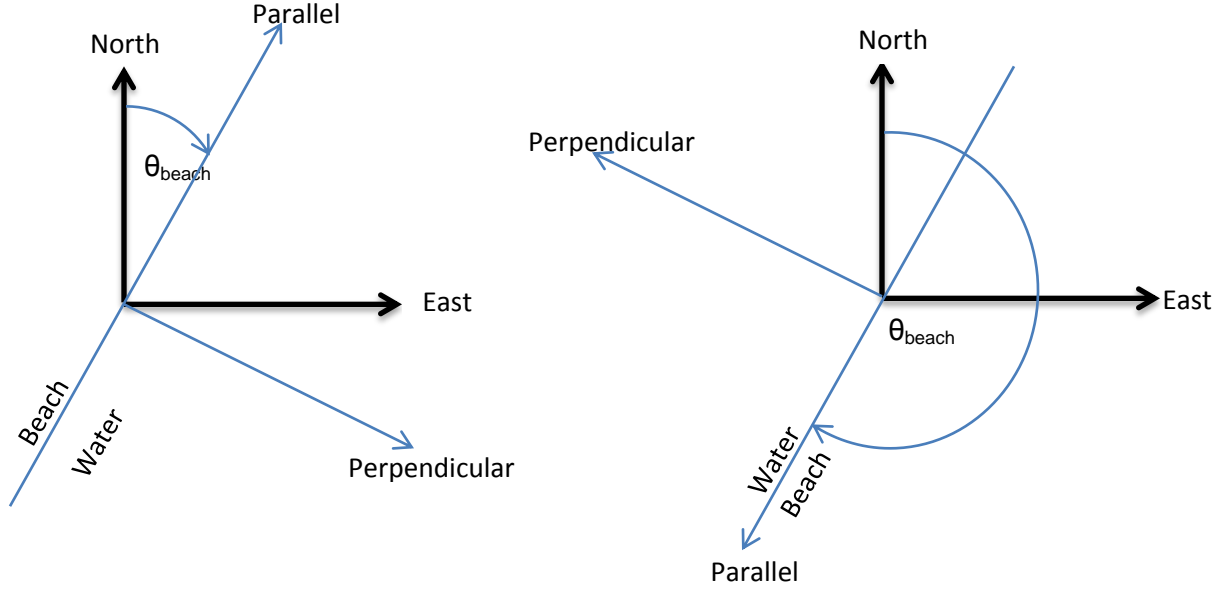
$$\sigma = \sqrt{\frac{1}{N} \sum_{i=0}^N (x_i - \mu)^2} \quad 6$$

A simple example is provided to demonstrate how N is found. The requested period is 60 minutes.

Time(minutes)	Value(x)	Mean( $\mu$ )	Min	Max	Sum( $\Sigma$ )	Diff( $\Delta$ )	StDev( $\sigma$ )
0	1						
15	3						
30	2						
45	1	1.75	1	3	7	0	0.829
60	1	1.75	1	3	7	-2	0.829
75	2	1.5	1	2	6	0	0.5
90	1	1.25	1	2	5	0	0.433

## B. Resolving Velocity Vectors

When velocity vectors are offered in east and north components, EnDDaT has the option to resolve those to parallel and perpendicular vectors based on a beach angle. The beach angle ( $\theta$ ) is defined as degrees off of north.



The basic method to convert north (N) and east(E) vectors to parallel ( $\parallel$ ) and perpendicular ( $\perp$ ) is to convert the N E vectors to polar coordinates, rotate the angle, and convert back to Cartesian coordinate. *mod* represents the modulo (remainder).

$$R = \sqrt{v_N^2 + v_E^2} \quad 7$$

$$\theta = \tan^{-1} \frac{v_E}{v_N} \quad 8$$

$$\theta_{new} = (\theta - \theta_{beach} + 2\pi) \bmod 2\pi \quad 9$$

$$v_{\parallel} = R \cos \theta_{new} \quad 10$$

$$v_{\perp} = R \sin \theta_{new} \quad 11$$

Converting the wave direction (W) to it's parallel and perpendicular unit vectors is similar:

$$\theta_{new} = (\theta_{wave\ direction} - \theta_{beach} + 2\pi) \bmod 2\pi \quad 12$$

$$W_{\parallel} = \cos \theta_{new} \quad 13$$

$$W_{\perp} = \sin \theta_{new} \quad 14$$

The wave direction unit vectors can be scaled with wave height( $H_{wave}$ ) as the magnitude:

$$H_{wave,||} = H_{wave}W_{||} \quad 15$$

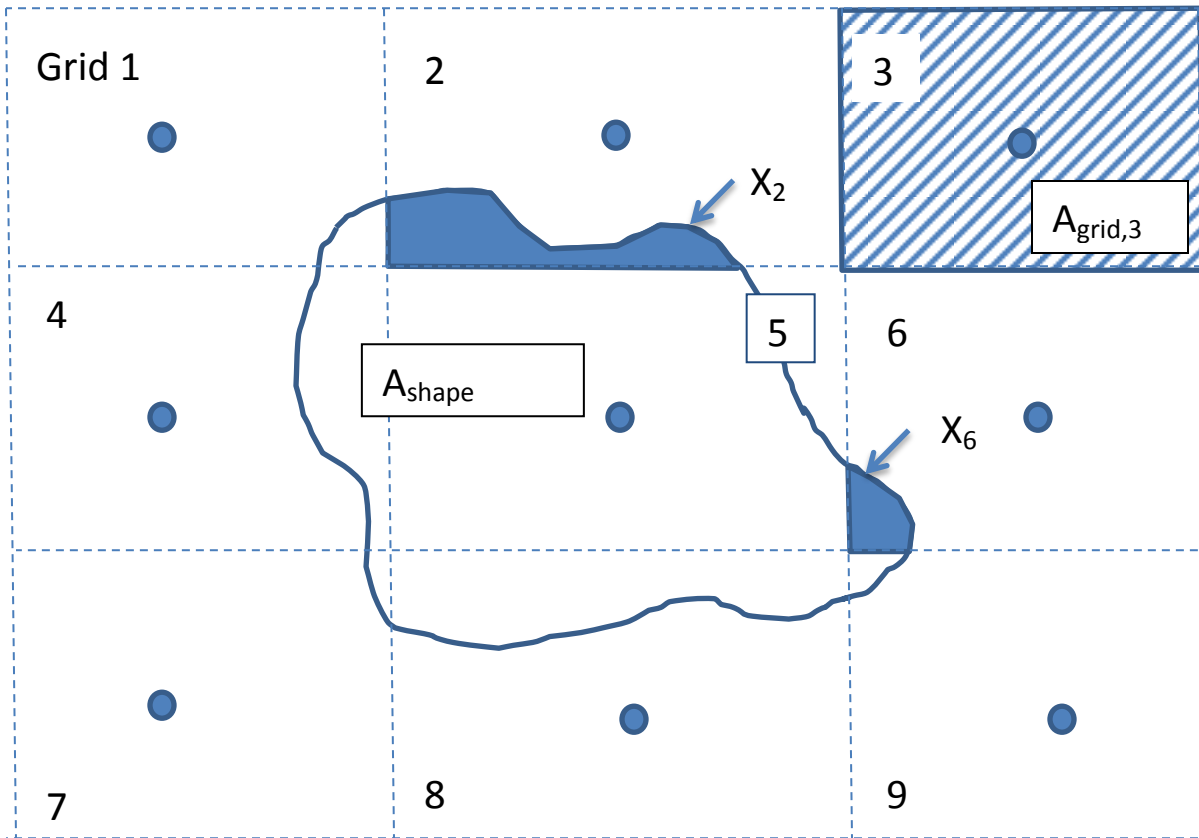
$$H_{wave,\perp} = H_{wave}W_{\perp} \quad 16$$

### C. Geographical processing

The algorithm used for the geoprocessing performed by the Geo Data Portal is described in the following reference:

D. H. D. West (1979). *Communications of the ACM*, 22, 9, 532-535: *Updating Mean and Variance Estimates: An Improved Method*

Briefly, the processes can be summarized as follows (although see above reference for complete details and caveats):



Assume that the blue dots represent the center of each grid, and the outline of each grid box is represented by the dashed blue line. Assume for this example that each grid box has an area  $A_{grid,i}$ . For each grid, there is a reported value. For this example, let's assume we are measuring precipitation, so we will call each grid's reported value  $P_i$ , where  $i$  is the grid number (1-9). The area of the shape that is in each grid quadrant is represented by  $X_i$ . The total area of the shape is  $A_{shape}$ . The following equations are used for the geometric processing (as defined by this example):

$$\text{Weighted Sum} = \sum_{i=1}^9 \frac{X_i}{A_{grid,i}}$$

17

$$\text{Weighted Mean} = \mu = \frac{\sum_{i=1}^9 P_i \frac{X_i}{A_{grid,i}}}{\sum_{i=1}^9 \frac{X_i}{A_{grid,i}}}$$

18

$$\text{Minimum} = \min(P_i) \{1, i, 9\}$$

19

$$\text{Maximum} = \max(P_i) \{1, i, 9\}$$

20

$$\text{Standard Deviation} = \sigma = \sqrt{\frac{1}{9-1} \sum_{i=1}^9 \left( P_i \frac{X_i}{A_{grid,i}} - \mu \right)^2}$$

21

22

$$\text{Variance} = \sigma^2$$

21

$$\text{Count} = \text{number of grid cells} = 9$$

23

## Appendix 3: Version Updates

### Beta Service Version 1.2.10

- May 18, 2012
- EnDDaT Introduction page added and set as opening page
- 'Home' page (functional data discovery page) set to /dataDiscovery.jsp
- Sortable table added to QW data discovery (see User Guide for more information)
- Cosmetic improvements to tables
- User Guide updated

### Beta Service Version 1.2.7

- Release April 30, 2012
- Data is offered by streaming service. This is much more robust than previous version
- Data can be directly downloaded in tab delimited or csv output
- Graphing outputs improved
- Data discovery streamlined

### Beta Service Version 1.2.3

- Release January 27, 2012
- Standard deviation and difference added to data processing options
- Included processing to convert wave direction or wave height to parallel and perpendicular components based on beach orientation.
- Added a hour/day/week toggle on the data processing page.
- Updated user guide to include more information about geometric processing.

### Beta Service Version 1.2.2:

- Released December 19, 2011:
- Version update page created (link found on Sidebar)
- Standard deviation and difference added to data processing options
- Beach orientation calculator moved to a separate tab on the Process Data page
- Found and fixed an indexing bug in the GLCFS data. Original data was off by 1 index (2 km)
- Data processing changed from 'inclusive' to 'exclusive'. For example:

**Old Method:** for a 1 hour calculation: 3 30-minute measurements are used

Time	Value	Mean over 1 hour	Summation over 1 hour
12:00	1		
12:30	2		
1:00	3	2	6

1:30	4	3	9
------	---	---	---

**New Method:** for a 1 hour calculation,: 2 30-minute measurements are used

Time	Value	Mean over 1 hour	Summation over 1 hour
12:00	1		
12:30	2	1.5	3
1:00	3	2.5	5
1:30	4	3.5	7

### Initial Release

- Version 0.0.0 - December 2, 2011